数理解析研究所講究録 1439 巻 2005 年 66-78

On the improved estimation of error variance and order restricted normal variances

Youhei Oono
School of Science for OPEN and Environmental Systems,
Graduate School of Science and Technology,
Keio University

Nobuo Shinozaki
Department of Administration Engineering,
Faculty of Science and Technology,
Keio University

Abstract

We consider the estimation of error variance and construct a class of estimators which uniformly improve upon the usual estimators. We also consider the estimation of order restricted normal variances. We give a class of isotonic regression estimators which uniformly improve upon the usual estimators including the unbiased estimator, the unrestricted maximum likelihood estimator and the best scale and translation equivariant estimator under various types of order restrictions. They are discussed under entropy loss and under squared error loss.

1. Introduction

Let S_0/σ^2 and S_i/σ^2, $i=1,2,\cdots,k$ be mutually independently distributed as $\chi^2_{\nu_0}$ and $\chi^2_{\nu_i}(\lambda_i)$, $i=1,2,\cdots,k$ respectively, where $\chi^2_{\nu_0}$ denotes the χ^2 distribution with ν_0 degrees of freedom and $\chi^2_{\nu_i}(\lambda_i)$ the noncentral χ^2 distribution with ν_i degrees of freedom and noncentrality parameter λ_i. Considering the estimation of variance σ^2 based on a random sample $X_1,\cdots,X_n$ from a normal population with unknown mean μ, it corresponds to the case when $k=1$, $S_0=\sum_{i=1}^{n}(X_i-\bar{X})^2$, $\nu_0=n-1$, $S_1=n\bar{X}^2$, $\nu_1=1$ and $\lambda_1=n\mu^2/(2\sigma^2)$. If we consider the estimation of error variance σ^2 based on experiments using two-level orthogonal arrays, S_0 and S_i are sum of squares for error term and that for each factorial effect, respectively.

When we estimate σ^2 under the squared error loss

$$L_1(\sigma^2, \hat{\sigma^2}) = \left(\hat{\sigma^2}/\sigma^2 - 1\right)^2, \tag{1}$$

the estimator $\delta_0 = S_0/(\nu_0 + 2)$ is the best among estimators of the form cS_0, where c is a constant. Stein (1964) showed that for the case when $k = 1$, $\delta_1 = \min\{S_0/(\nu_0 + 2), (S_0 + S_1)/(\nu_0 + \nu_1 + 2)\}$ uniformly improves upon δ_0. Gelfand and Dey (1988) generalized Stein's result and showed that

$$\delta_0 \prec \delta_1 \prec \cdots \prec \delta_k, \tag{2}$$

where δ_j is the estimator defined by $\delta_j = \min_{0\le l\le j}[(\sum_{i=0}^{l} S_i)/(\sum_{i=0}^{l} \nu_i + 2)]$, $j = 1, \cdots, k$ and $\delta_j \prec \delta_{j+1}$ means that δ_{j+1} uniformly improves upon δ_j. One may think that it is more appropriate to consider the estimation of σ^2 under the entropy loss function

$$L_2(\sigma^2, \hat{\sigma^2}) = \hat{\sigma^2}/\sigma^2 - \log(\hat{\sigma^2}/\sigma^2) - 1. \tag{3}$$

Then, it is well-known that the best positive multiple of S_0 is the unbiased estimator

$$\zeta_0 = S_0/\nu_0, \tag{4}$$

and that it is improved upon uniformly by a Stein-type shrinkage estimator when $k = 1$. (See Brown (1968) and Brewster and Zidek (1974).)

In Section 2, we first construct a wide class of estimators of σ^2, which uniformly improve upon the positive multiples of S_0 under the entropy loss (3). Further, under the squared error loss (1), we construct a class of improved estimators of σ^2, which gives a generalization of the result (2).

These results are applied to the estimation problem of order restricted normal variances. Let X_{ij} be the j-th observation from the i-th population and be mutually independently distributed as $N(\mu_i, \sigma_i^2)$, $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, n_i$, where μ_i's are unknown. Let us define $V_i = \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$, then V_i's are mutually independently distributed as $\sigma_i^2\chi^2_{\nu_i}$, where $\nu_i = n_i - 1$. Assume that it is known that

(A.1) σ_1^2 is the smallest among σ_i^2, $i = 1, 2, \cdots, k$.

When we estimate σ_1^2 assuming the simple order restriction $\sigma_1^2 \le \cdots \le \sigma_k^2$, the isotonic regression estimator based on V_i/ν_i with weights ν_i is given by

$$\tilde{\sigma_1^2}^{SO} = \min_{1\le j\le k}[(\sum_{l=1}^{j} \nu_l(V_l/\nu_l))/(\sum_{l=1}^{j} \nu_l)]. \tag{5}$$

Hwang and Peddada (1994) showed that when it is known that (A.1), $\tilde{\sigma}_1^{2\,SO}$ uniformly improves upon V_1/ν_1 under the loss function $L(\sigma_1^2, \hat{\sigma}_1^2) = \rho(|\hat{\sigma}_1^2 - \sigma_1^2|)$, where $\rho(\cdot)$ is an arbitrary nondecreasing function. (Regarding this loss, see Hwang (1985).)

In Section 3, for the case when it is known that (A.1), we first construct a class of estimators based on V_i's which uniformly improve upon usual estimators of σ_1^2 including the unbiased estimator, the unrestricted maximum likelihood estimator and the best scale and translation equivariant estimator. They are considered under entropy loss and under squared error loss. Our improved estimator is considered as isotonic regression estimator under dummy simple order restriction. Further, we mention that the results can be applied to the estimation of each variance under various order restrictions. Finally, we show that our improved estimator can be further improved upon uniformly by an estimator using not only V_i's but also $\bar{X}_i$'s.

2. A class of improved estimators of variance

Let S_0 and S_i, $i = 1, 2, \cdots, k$ be random variables distributed as stated in the Introduction. We construct a class of estimators of σ^2 improving upon the positive multiple of S_0 directly under the entropy loss (3) and also under the squared error loss (1).

2.1 Improved estimators under entropy loss

To give a class of improved estimators under entropy loss, we first show Theorem 2.1 using the following Lemma, which was given in Shinozaki (1995).

Lemma 2.1. For $0 \le v < 1$,

$$\log(1 - v) \ge -v - \frac{v^2}{6} - \frac{v^2}{3(1 - v)}.$$

Theorem 2.1. For $1 \le j \le k$, let $\phi_j : \mathbb{R}^j \to \mathbb{R}^1$ be positive real valued function of

$$\gamma_j = \left(\frac{S_0}{S_0 + S_1}, \frac{S_0 + S_1}{S_0 + S_1 + S_2}, \cdots, \frac{\sum_{i=0}^{j-1} S_i}{\sum_{i=0}^{j} S_i} \right),$$

and let $a_j \ge 1/(\sum_{i=0}^{j} \nu_i)$. When we estimate σ^2 under entropy loss, $\min\{\phi_j(\gamma_j), a_j\} \sum_{i=0}^{j} S_i$ uniformly improves upon $\phi_j(\gamma_j) \sum_{i=0}^{j} S_i$ if $\phi_j(\gamma_j) > a_j$ with positive probability.

Proof. Let us denote $\tilde{\sigma^2} = \phi_j(\gamma_j)\sum_{i=0}^{j} S_i$ and $\hat{\sigma^2} = \min\{\phi_j(\gamma_j), a_j\}\sum_{i=0}^{j} S_i$. Noting that $\hat{\sigma^2}$ can be expressed as

$$\hat{\sigma^2} = (\sum_{i=0}^{j} S_i)\phi_j(\gamma_j) - (\sum_{i=0}^{j} S_i)(\phi_j(\gamma_j) - a_j)I_{\phi_j(\gamma_j)\geq a_j}, \tag{6}$$

where I_C denotes the indicator function of the set satisfying the condition C, we have the loss difference of $\tilde{\sigma^2}$ and $\hat{\sigma^2}$ as

$$\begin{aligned} &L_2(\sigma^2, \tilde{\sigma^2}) - L_2(\sigma^2, \hat{\sigma^2}) \\ &= \left(\frac{\sum_{i=0}^{j} S_i}{\sigma^2}\right)(\phi_j(\gamma_j) - a_j)I_{\phi_j(\gamma_j)\geq a_j} + \log\left\{1 - \left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right) I_{\phi_j(\gamma_j)\geq a_j}\right\}. \end{aligned} \tag{7}$$

Noting that $0 \leq \{1 - a_j/\phi_j(\gamma_j)\}I_{\phi_j(\gamma_j)\geq a_j} < 1$ and using Lemma 2.1, we evaluate the second term on the right-hand side of (7) as

$$\begin{aligned} &\log\left\{1 - \left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right) I_{\phi_j(\gamma_j)\geq a_j}\right\} \\ &\geq -\left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right) I_{\phi_j(\gamma_j)\geq a_j} - \frac{1}{6}\left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right)^2 I_{\phi_j(\gamma_j)\geq a_j} \\ &\quad - \frac{1}{3}\frac{\left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right)^2 I_{\phi_j(\gamma_j)\geq a_j}}{1 - \left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right) I_{\phi_j(\gamma_j)\geq a_j}} \\ &= \left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right)\frac{\phi_j(\gamma_j)}{a_j}\left\{\frac{1}{6}\left(\frac{a_j}{\phi_j(\gamma_j)}\right)^2 - \frac{5}{6}\frac{a_j}{\phi_j(\gamma_j)} - \frac{1}{3}\right\} I_{\phi_j(\gamma_j)\geq a_j}, \end{aligned} \tag{8}$$

where the last equality is by

$$\frac{\left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right)^2 I_{\phi_j(\gamma_j)\geq a_j}}{1 - \left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right) I_{\phi_j(\gamma_j)\geq a_j}} = \frac{\phi_j(\gamma_j)}{a_j}\left(1 - \frac{a_j}{\phi_j(\gamma_j)}\right)^2 I_{\phi_j(\gamma_j)\geq a_j}. \tag{9}$$

To evaluate the expectation of (7), we introduce auxiliary random variables K_i, $i = 1, \cdots, j$ distributed independently as Poisson distribution with mean λ_i such that K_i is independent of S_0, and S_i given K_i is distributed as $\sigma^2\chi^2_{\nu_i+2K_i}$. Note that given $K = (K_1, \cdots, K_j)$, $\sum_{i=0}^{j} S_i$ and γ_j are mutually independent and that $\sum_{i=0}^{j} S_i$ given K is distributed as $\sigma^2\chi^2_{\nu_0+\sum_{i=1}^{j}(\nu_i+2K_i)}$. Thus we evaluate the expectation of the first term on the right-hand side of (7) given K as

$$E\left[\left(\frac{\sum_{i=0}^{j} S_i}{\sigma^2}\right)(\phi_j(\gamma_j)-a_j)I_{\phi_j(\gamma_j)\geq a_j} \mid K\right]$$
$$= a_j\left\{\nu_0+\sum_{i=1}^{j}(\nu_i+2K_i)\right\}E\left[\left(1-\frac{a_j}{\phi_j(\gamma_j)}\right)\frac{\phi_j(\gamma_j)}{a_j}I_{\phi_j(\gamma_j)\geq a_j} \mid K\right]$$
$$\geq E\left[\left(1-\frac{a_j}{\phi_j(\gamma_j)}\right)\frac{\phi_j(\gamma_j)}{a_j}I_{\phi_j(\gamma_j)\geq a_j} \mid K\right], \tag{10}$$

where we have the last inequality from $a_j \geq 1/(\sum_{i=0}^{j}\nu_i)$. Using (8) and (10), we see that the expectation of (7) given K is not smaller than

$$\frac{1}{6}E\left[\frac{\phi_j(\gamma_j)}{a_j}\left(1-\frac{a_j}{\phi_j(\gamma_j)}\right)\left\{\left(\frac{a_j}{\phi_j(\gamma_j)}\right)^2-5\frac{a_j}{\phi_j(\gamma_j)}+4\right\}I_{\phi_j(\gamma_j)\geq a_j} \mid K\right]$$
$$=\frac{1}{6}E\left[\frac{\phi_j(\gamma_j)}{a_j}\left(1-\frac{a_j}{\phi_j(\gamma_j)}\right)^2\left(4-\frac{a_j}{\phi_j(\gamma_j)}\right)I_{\phi_j(\gamma_j)\geq a_j} \mid K\right], \tag{11}$$

which is clearly positive since $\phi_j(\gamma_j) > a_j$ with positive probability. Taking the expectation of (11) over K, we see that the risk of $\hat{\sigma^2}$ is smaller than that of $\tilde{\sigma^2}$ and this completes the proof. □

Based on Theorem 2.1, we construct a class of estimators improving upon estimators of the form

$$\eta_0 = a_0 S_0, \tag{12}$$

where a_0 is a positive constant. The estimator ζ_0 is clearly of the form (12). Though an estimator improving upon the best positive multiple ζ_0, uniformly improves upon η_0, we are also interested in constructing a class of estimators improving upon η_0 directly. We first note that η_0 can be written as $\eta_0 = \phi_1(\gamma_1)(S_0+S_1)$, where $\phi_1(\gamma_1) = a_0\gamma_1$ and $\gamma_1 = S_0/(S_0+S_1)$. Let

$$\eta_j = \phi_{j+1}(\gamma_{j+1})\sum_{i=0}^{j+1} S_i, \tag{13}$$

with

$$\phi_{j+1}(\gamma_{j+1}) = \min\{\phi_j(\gamma_j), a_j\}\left(\frac{\sum_{i=0}^{j} S_i}{\sum_{i=0}^{j+1} S_i}\right) \tag{14}$$

for $j = 1, 2, \cdots, k-1$ and let

$$\eta_k = \min\{\phi_k(\gamma_k), a_k\}\sum_{i=0}^{k} S_i. \tag{15}$$

(Note that the right-hand side of (14) is a function of γ_{j+1}.) Then η_{j-1} and η_j can be expressed as $\phi_j(\gamma_j)\sum_{i=0}^{j} S_i$ and $\min\{\phi_j(\gamma_j), a_j\}\sum_{i=0}^{j} S_i$, respectively. Thus from Theorem 2.1, we see that η_j uniformly improves upon η_{j-1} if $a_j \geq 1/(\sum_{i=0}^{j}\nu_i)$ and $a_{i-1} > a_j$, $i = 1, \cdots, j$. Using (12), (13), (14) and (15) inductively, we see that η_j is also expressed as $\min_{0\leq l\leq j}[a_l(\sum_{i=0}^{l} S_i)]$, and we have the following Theorem.

Theorem 2.2. Let $a_0 > 1/(\nu_0 + \nu_1)$ and let $\eta_j = \min_{0\leq l\leq j}[a_l(\sum_{i=0}^{l} S_i)]$, $j = 0, 1, \cdots, k$. Under entropy loss,

$$\eta_0 \prec \eta_1 \prec \cdots \prec \eta_k, \tag{16}$$

if $a_j \geq 1/(\sum_{i=0}^{j}\nu_i)$ and $a_{j-1} > a_j$, $j = 1, 2, \cdots, k$.

From Theorem 2.2, we see that η_j, $j = 1, 2, \cdots, k$ constitute a class of estimators which uniformly improve upon η_0. We should remark that this class is determined by a_j, $j = 1, \cdots, k$.

Remark 2.1. For fixed a_0, we can choose specific values of $a_1, \cdots, a_k$ satisfying the condition given in Theorem 2.2. One such choice is $a_j = 1/(\sum_{i=0}^{j}\nu_i)$, $j = 1, \cdots, k$ for $a_0 = 1/\nu_0$ and under entropy loss we have

$$\zeta_0 \prec \zeta_1 \prec \cdots \prec \zeta_k, \tag{17}$$

where ζ_0 is as defined by (4) and

$$\zeta_j = \min_{0\leq l\leq j}[(\sum_{i=0}^{l} S_i)/(\sum_{i=0}^{l}\nu_i)],\ j = 1, 2, \cdots, k. \tag{18}$$

Note that ζ_0 is the best estimator of the form (12) under entropy loss as well as the unbiased estimator.

2.2 Improved estimators under squared error loss

Here, under the squared error loss (1), we give a class of improved estimators of σ^2, which are slight modifications of the estimators given by Gelfand and Dey (1988). They are given in the following Theorem, whose proof is similar to that of Theorem 1 in Gelfand and Dey (1988) and is omitted here.

Theorem 2.3. Let $a_0 > 1/(\nu_0 + \nu_1 + 2)$ and let $\eta_j = \min_{0\leq l\leq j}[a_l(\sum_{i=0}^{l} S_i)]$, $j = 0, 1, \cdots, k$. Under squared error loss,

$$\eta_0 \prec \eta_1 \prec \cdots \prec \eta_k, \tag{19}$$

if $a_j \geq 1/(\sum_{i=0}^{j} \nu_i + 2)$ and $a_{j-1} > a_j$, $j = 1, 2, \cdots, k$.

Remark 2.2. For fixed a_0, we can choose specific values of $a_1, \cdots, a_k$ satisfying the conditions given in Theorem 2.3. One such choice is (a) $a_j = 1/(\sum_{i=0}^{j} \nu_i + 2)$, $j = 1, \cdots, k$ for $a_0 = 1/(\nu_0 + 2)$ and we have (2) which is given by Gelfand and Dey (1988). Another choice is (b) $a_j = 1/(\sum_{i=0}^{j} \nu_i)$, $j = 1, \cdots, k$ for $a_0 = 1/\nu_0$ and we have (17) under squared error loss, which constitutes a class of improved estimators over the unbiased estimator S_0/ν_0. We note that Nagata (1989) has given the estimator for the case when $k = 1$ essentially.

3. An application to the estimation problem of ordered variances

In this section, under entropy loss and under squared error loss, we discuss the estimation of order restricted normal variances. Let X_{ij}, $i = 1, 2, \cdots, k$, $j = 1, 2, \cdots, n_i$ be the j-th observation of the i-th population and be mutually independently distributed as $N(\mu_i, \sigma_i^2)$, where μ_i's are unknown. Let us define $V_i = \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$, then V_i's are mutually independently distributed as $\sigma_i^2 \chi_{\nu_i}^2$, where $\nu_i = n_i - 1$. Assume that it is known that (A.1).

3.1 Improved estimation of each variance

We first consider the improved estimation of σ_1^2 based on V_i, $i = 1, 2, \cdots, k$. Note that $V_1/(\nu_1 + 1)$ is the unrestricted maximum likelihood estimator and V_1/ν_1 (or $V_1/(\nu_1 + 2)$) is the best scale and translation equivariant estimator under entropy loss (or under squared error loss). In the following, we construct a class of estimators, which uniformly improve upon usual estimators of the form cV_1. The following well-known Lemma is a preliminary for our discussion.

Lemma 3.1. Let V_i be distributed as $\sigma_i^2 \chi_{\nu_i}^2$, where $\sigma_i^2 \geq \sigma_1^2$. Then there exists an auxiliary random variable U_i satisfying the following two conditions.
(a) V_i given U_i is distributed as $\sigma_1^2 \chi_{\nu_i}^2(U_i)$.
(b) U_i is distributed as $\tau_i^2/(2\sigma_1^2)\chi_{\nu_i}^2$, where $\tau_i^2 = \sigma_i^2 - \sigma_1^2$.

Now, based on the results of Theorems 2.2 and 2.3 and Lemma 3.1, we show that the estimator

$$\hat{\sigma_1^2}^S = \min_{1 \leq j \leq k} [(\sum_{l=1}^{j} V_l)/(\sum_{l=1}^{j} w_l)] \tag{20}$$

uniformly improves upon V_1/w_1 if the weights w_i, $i = 1, \cdots, k$ satisfy some conditions, which we state in the following Theorem.

Theorem 3.1. Assume that it is known that σ_1^2 is the smallest among σ_i^2's.
(i) Let $0 < w_1 < \nu_1+\nu_2$. Under entropy loss, the estimator $\hat{\sigma_1^2}^S$ uniformly improves upon V_1/w_1 if $w_2, \cdots, w_k$ satisfy $\sum_{l=1}^{j} w_l \leq \sum_{l=1}^{j} \nu_l$ and $w_j > 0$, $j = 2, \cdots, k$.
(ii) Let $0 < w_1 < \nu_1 + \nu_2 + 2$. Under squared error loss, the estimator $\hat{\sigma_1^2}^S$ uniformly improves upon V_1/w_1 if $w_2, \cdots, w_k$ satisfy $\sum_{l=1}^{j} w_l \leq \sum_{l=1}^{j} \nu_l + 2$ and $w_j > 0$, $j = 2, \cdots, k$.
Proof. We only deal with (i) since (ii) can be proved similarly. From Lemma 3.1, we can imagine auxiliary independent random variables U_i, $i = 2, \cdots, k$ such that V_1 and V_i, $i = 2, \cdots, k$ given U_i, $i = 2, \cdots, k$ are mutually independently distributed as $\sigma_1^2\chi_{\nu_1}^2$ and $\sigma_1^2\chi_{\nu_i}^2(U_i)$, $i = 2, \cdots, k$ respectively. Given U_i, $i = 2, \cdots, k$, by applying Theorem 2.2 with $S_i = V_{i+1}$, $i = 0, 1, \cdots, k-1$, $\nu_i = \nu_{i+1}$, $i = 0, 1, \cdots, k-1$, $\lambda_i = U_{i+1}$, $i = 1, 2, \cdots, k-1$ and $a_i = 1/(\sum_{l=1}^{i+1} w_l)$, $i = 0, 1, \cdots, k-1$, we have $\eta_0 \prec \eta_{k-1}$, which is equivalent to

$$E[L_1(\sigma_1^2, \hat{\sigma_1^2}^S)|U_2, \cdots, U_k] < E[L_1(\sigma_1^2, V_1/w_1)|U_2, \cdots, U_k]. \tag{21}$$

Taking the expectation on both sides of (21) over $U_2, \cdots, U_k$, we see that (i) is true and this completes the proof. □

Note. We should mention that (ii) of Theorem 3.1 gives a generalization of Theorem 2 in Gelfand and Dey (1988) who also utilized our Lemma 3.1 in their proof.

Remark 3.1 For fixed w_1, we can choose specific values of weights $w_2, \cdots, w_k$ satisfying the conditions given in Theorem 3.1 and we have estimators improving upon the unrestricted maximum likelihood estimator, the unbiased estimator and the best scale and translation equivariant estimator. For example: (a) If we choose $w_i = \nu_i$, $i = 2, \cdots, k$ for $w_1 = \nu_1$ in (i), we see that under entropy loss the estimator (5) uniformly improves upon the best scale and translation equivariant estimator V_1/ν_1. (b) If we choose $w_i = \nu_i$, $i = 2, \cdots, k$ for $w_1 = \nu_1$ in (ii), we see that the estimator (5) uniformly improves upon the unbiased estimator V_1/ν_1, which is the result implied by Hwang and Peddada (1994) under squared error loss. (c) If we choose $w_2 = \nu_2 - 1$ and $w_i = \nu_i$, $i = 3, \cdots, k$ for $w_1 = \nu_1 + 1$ in (i) and (ii), we have an estimator improving upon the unrestricted maximum likelihood estimator for both loss functions. (Note that in case of (c), we assume that $\nu_2 \geq 2$.)

Remark 3.2. Since the estimator $\hat{\sigma_1^2}^S$ can be written as

$$\hat{\sigma_1^2}^S = \min_{1 \leq j \leq k} [\{\sum_{l=1}^{j} w_l(V_l/w_l)\}/(\sum_{l=1}^{j} w_l)], \tag{22}$$

it can be considered as the isotonic regression estimator of σ_1^2 based on V_i/w_i with weights w_i under the simple order restriction $\sigma_1^2 \leq \cdots \leq \sigma_k^2$. (See Robertson, Wright and Dykstra (1988) or Barlow, Bartholomew, Bremner and Brunk (1972).) Note that this estimator is not the isotonic regression when it is known that (A.1). In this remark, without loss of generality, we assume that $\sigma_i^2 \leq \sigma_j^2$ if the ordering between σ_i^2 and σ_j^2, $2 \leq i < j \leq k$ is known. Then Theorem 3.1 implies the following about this estimator. The ordering between $\sigma_2^2, \cdots, \sigma_k^2$ is not completely known, so we guess it, while preserving the known ordering, and construct dummy simple order restriction: $\sigma_1^2 \leq \cdots \leq \sigma_k^2$. Theorem 3.1 assures that the isotonic regression estimator under this dummy simple order restriction uniformly improves upon V_1/w_1 even if the guess is wrong. Note that w_i's must satisfy the conditions given in Theorem 3.1.

Theorem 3.1 can be applied to the estimation of each variance under various types of order restrictions. Before proceeding any further, we introduce a pictorial notation of order restriction developed by Hwang and Peddada (1994). In Fig. 1, each graph ((a)-(d)) represents the corresponding order restriction. For example Fig. 1 (a) corresponds to the simple order restriction $\sigma_1^2 \leq \sigma_2^2 \leq \sigma_3^2 \leq \sigma_4^2$. Note that σ_i^2's are denoted by solid circles. We omit writing σ^2 on the graphs but only write the subscripts. If two circles are joined together by a line segment, it means that the circle with larger number is known to correspond to the larger σ^2. For example Fig. 1 (b) corresponds to the order restriction $\sigma_1^2 \leq \sigma_2^2, \sigma_3^2 \leq \sigma_4^2 \leq \sigma_5^2, \sigma_6^2 \leq \sigma_7^2$.

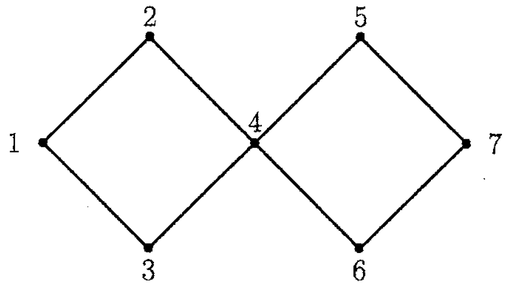
Now, we explain an improved estimation scheme. We should mention that Hwang and Peddada (1994) proposed similar procedure for estimating order restricted location parameters of elliptically symmetric distributions. We first consider the case when it is known which variance corresponds to the smallest variance (e.g. Fig 1 (a) and (b)). Without loss of generality, we assume that σ_1^2 is the smallest variance. The estimation procedure is given as follows.

Step 1. Estimation of σ_1^2. From Theorem 3.1 and Remark 3.2, we can construct an isotonic regression estimator of σ_1^2 which gives the uniform improvement over V_1/w_1 if w_1 is not so large as shown in Theorem 3.1.

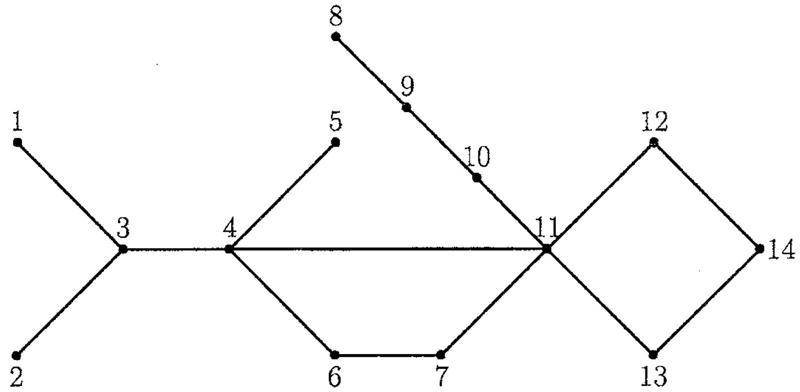
Step 2. Estimation of other variances. When we estimate σ_i^2, we remove the smallest number of circles from the graph so that σ_i^2 becomes the smallest variance in the resulting subgraph G_i. Then by Theorem 3.1, we can construct an isotonic regression estimator of σ_i^2 based on the circles in G_i, which gives the uniform improvement over V_i/w_i if w_i satisfies the condition implied by Theorem 3.1.

Example. When we consider the estimation of σ_3^2 in Fig 1 (b), we remove the circles 1 and 2 so that σ_3^2 corresponds to the smallest variance in the resulting subgraph Fig 1 (d). We guess the unknown ordering between σ_5^2 and σ_6^2 in the subgraph G_3, and we have the dummy simple order restriction $\sigma_3^2 \leq \sigma_4^2 \leq \sigma_5^2 \leq \sigma_6^2 \leq$

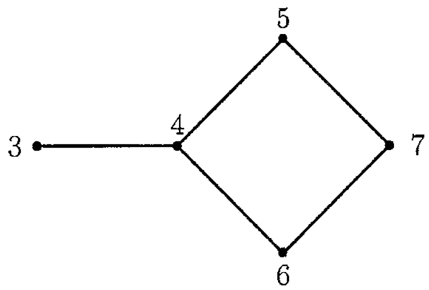
(a) Graph corresponding to $\sigma_1^2 \le \sigma_2^2 \le \sigma_3^2 \le \sigma_4^2$.



(b) Graph corresponding to $\sigma_1^2 \le \sigma_2^2, \sigma_3^2 \le \sigma_4^2 \le \sigma_5^2, \sigma_6^2 \le \sigma_7^2$.



(c) Graph when it is not known which variance is the smallest.



(d) Subgraph G_3 of (b).

Fig. 1. Pictorial representation of order restriction.

σ_7^2. Then under this dummy order restriction, we construct isotonic regression estimator of σ_3^2 based on $V_i/w_i,\ i = 3, 4, \cdots, 7$ with weights $w_i,\ i = 3, 4, \cdots, 7$, that is $\hat{\sigma_3^2}^S = \min_{3\le j\le k}[(\sum_{l=3}^{j} V_l)/(\sum_{l=3}^{j} w_l)]$, which gives the uniform improvement over V_3/w_3 if $w_i,\ i = 3, \cdots, 7$ satisfy some conditions. As for the estimation of $\sigma_2^2, \sigma_4^2, \sigma_5^2$ and σ_6^2, we can discuss similarly. However, our procedure does not work for the estimation of σ_7^2, the largest variance.

When it is not known which variance corresponds to the smallest variance (e.g. Fig 1 (c)), we can start with Step 2. We should notice here that though our scheme gives improved estimators of each of order restricted variances, the obtained estimates may violate the known order restriction unfortunately. To the best of our knowledge, it is not well established when and how we can construct such estimators which not only improve upon usual estimators but also preserve the known order restriction.

3.2 Further improvement

Here, we show that our improved estimator given in Section 3.1 can be further improved upon uniformly by an estimator which use not only V_i's but also $\bar{X}_i$'s. We give an estimator improving upon $\hat{\sigma_1^2}^S$ especially for the case when $k = 2$ and $\sigma_1^2 \le \sigma_2^2$ is known. We can similarly discuss the estimation of each of order restricted variances also for the case when $k \ge 3$. Let $Q_j = n_j \bar{X}_j^2,\ j = 1, 2$, then Q_j's are independently distributed as $\sigma_j^2 \chi_1^2(\lambda_j)$, where $\lambda_j = n_j \mu_j^2/(2\sigma_j^2)$. We can imagine random variables $K_j,\ j = 1, 2$ distributed independently as Poisson distributions with means $\lambda_j,\ j = 1, 2$ such that given K_j's, Q_j's are independently distributed as $\sigma_j^2 \chi_{1+2K_j}^2$ respectively. Further from Lemma 3.1, we can imagine a random variable T_2 such that T_2 given K_2 is distributed as $(\sigma_2^2 - \sigma_1^2)/(2\sigma_1^2)\chi_{1+2K_2}^2$ and that Q_2 given K_2 and T_2 is distributed as $\sigma_1^2 \chi_{1+2K_2}^2(T_2)$. Thus, together with the proof of Theorem 3.1, we can imagine auxiliary random variables U_2, K_1, K_2 and T_2 such that $V_1,\ V_2,\ Q_1$ and Q_2 given them are independently distributed as $\sigma_1^2\chi_{\nu_1}^2,\ \sigma_1^2\chi_{\nu_2}^2(U_2),\ \sigma_1^2\chi_{1+2K_1}^2$ and $\sigma_1^2\chi_{1+2K_2}^2(T_2)$. Note that $\hat{\sigma_1^2}^S$ is expressed as

$$\min\{a_1 V_1, a_2(V_1 + V_2)\}, \tag{23}$$

where a_1 and a_2 are given constants. Also note that when we consider the estimation of σ^2 under entropy loss (or squared error loss), a_1 and a_2 must satisfy the condition $a_1 > a_2 \ge 1/(\nu_1 + \nu_2)$ (or $a_1 > a_2 \ge 1/(\nu_1 + \nu_2 + 2)$). Similarly with the proof of Theorem 3.1, we see that $\hat{\sigma_1^2}^S$ is improved upon uniformly by

$$\min\{a_1 V_1, a_2(V_1 + V_2), a_3(V_1 + V_2 + Q_1), a_4(V_1 + V_2 + Q_1 + Q_2)\} \tag{24}$$

if $a_j \ge 1/(\nu_1 + \nu_2 + j - 2)$ and $a_{j-1} > a_j,\ j = 3, 4$ (or if $a_j \ge 1/(\nu_1 + \nu_2 + j)$ and $a_{j-1} > a_j,\ j = 3, 4$) under entropy loss (or under squared error loss).

We should mention that we can construct an estimator improving upon a_1V_1 by using V_1, V_2, Q_1 and Q_2 regardless of the pooling order of V_2, Q_1 and Q_2. For example,

$$\min\{a_1V_1, b_2(V_1+Q_1), b_3(V_1+Q_1+V_2), b_4(V_1+Q_1+V_2+Q_2)\} \tag{25}$$

and

$$\min\{a_1V_1, c_2(V_1+Q_2), c_3(V_1+Q_2+Q_1), c_4(V_1+Q_2+Q_1+V_2)\} \tag{26}$$

uniformly improve upon a_1V_1 if a_1, b_j, $j=2,3,4$ and c_j, $j=2,3,4$ satisfy some conditions which will be apparent from Theorems 2.2 and 2.3.

References

[1] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression. Wiley, New York.

[2] Brewster, J. F. and Zidek, J. V. (1974). Improving on equivariant estimators. Ann. Statist. 2, 21-38.

[3] Brown, L. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. Ann. Math. Statist. 39, 29-48.

[4] Gelfand, A. E. and Dey, D. K. (1988). Improved estimation of the disturbance variance in a linear regression model. J. Econometrics 39, 387-395.

[5] Hwang, J. T. (1985). Universal domination and stochastic domination: estimation simultaneously under a broad class of loss functions. Ann. Statist. 13, 295-314.

[6] Hwang, J. T. G. and Peddada, S. D. (1994). Confidence interval estimation subject to order restrictions. Ann. Statist. 22, 67-93.

[7] Nagata, Y. (1989). Estimation with pooling procedures of error variance in ANOVA for orthogonal arrays. J. Jap. Soc. Quality Control 19, 12-19. (in Japanese.)

[8] Oono, Y. and Shinozaki, N. (2004). Estimation of error variance in the analysis of experiments using two-level orthogonal arrays. Comm. Statist. Theory Methods 33, 75-98.

[9] Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). Order Restricted Statistical Inference. Wiley, New York.

[10] Shinozaki, N. (1995). Some modifications of improved estimators of a normal variance. Ann. Inst. Statist. Math. 47, 273-286.

[11] Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. Ann. Inst. Statist. Math. 16, 155-160.